Simulation of 8-Hour Ozone Concentrations for the State of Arizona

Contract Report Submitted to the
Arizona Department of Environmental Quality

By

S.M. Lee[*]
S. Grossman-Clarke[†]
H.J.S. Fernando[†]

Environmental Fluid Dynamics Program
[*]Department of Civil & Environmental Engineering
[†]Department of Mechanical & Aerospace Engineering

1. Introduction

On November 14, 2002, the USEPA issued a memorandum entitled "Schedule for 8-Hour Ozone Designations and its Effect on Early Action Compacts," wherein the states and tribes were requested to provide recommendations for 8-hour ozone designations no later than April 15,[th] 2003. This was to include specific boundaries of the proposed non-attainment areas, supporting (2001-2003) air quality data and any other documentation relevant to the states' designations. In a follow-up memo by the EPA dated February 27, 2003, an extension was granted until July 15, 2003, to comply with this request, and the date of promulgation of non-compliance areas has been set for April 15, 2004.

In support of their efforts in meeting the EPA guidelines, the Arizona Department of Environmental Quality (ADEQ) contracted the Environmental Fluid Dynamics (EFD) Program at ASU to undertake the following tasks: (i) develop an emission inventory by combining the Maricopa Association of Governments (MAG) emissions inventory with the Western Region Air Partnership (WRAP) inventory to expand the present MAG modeling domain while including such population centers as Casa Grande, Coolidge and Florence; (ii) conduct 8-hour ozone simulations using Models-3/CMAQ and MM5 meteorological models for one or two design dates (in 2001 or 2002) selected by the ADEQ staff; this selection is to be based on the zone concentrations on the eastern fringe of the metro area and the availability of measurements from the far eastern part of the valley, such as Queen Valley and Tonto National Monument; (iii) validate the model's ozone output against the observations; (iv) develop a 2018 emission inventory, similar in nature to the CMAQ emissions for WRAP, by considering growth and control factors; (iv) conduct a socio-economic analysis based on GIS techniques that will outline growth scenarios for the greater Phoenix area.

Some of the work listed above was subcontracted to carefully selected experts in air quality modeling and analysis, from within and outside ASU. The GIS laboratory at ASU performed the socioeconomic analysis, and Mark Houyoux of the North Carolina Super Computer Center, Environmental programs, compiled the 2018 pollution inventory. These groups will submit separate reports to ADEQ. The present report contains the building of the pollution inventory beyond the MAG domain, Models-3/CMAQ simulations and the validation of meteorological and ozone modeling conducted by the EFD Program.

2. Design Days For Numerical Simulations

ADEQ recommended two design days for simulations based on the observations of elevated 8-hour ozone concentrations. The first is June 6, 2002 wherein high ozone concentrations were measured in the northeast part of the valley. The second day is July 12, 2002, where elevated 8-hour concentrations were recorded at Humboldt Mountain and in the central valley area.

On June 6, 2002, 8-hour ozone concentrations at Fountain Hills, Blue Point Bridge, Rio Verde and Tonto National Monument were respectively, 93, 92, 90, and 89 ppb.

During this episode, hot and clear weather was observed due to a high pressure system located over Arizona and a thermal low was found to form over the arid area in the vicinity of the Arizona, California and Mexico border. A meteorological condition with light surface wind and strong shortwave radiation was favorable for photochemical production of ozone and for the transport of a high-ozone laden air mass to far downwind of the valley.

The highest 8-hour ozone concentration on July 12th was measured at the Humboldt Mountain (103 ppb), and the next highest concentrations of 94, 93, 90, and 89 ppb were recorded at Falcon Field, Fountain Hills, Blue Point Bridge, and North Phoenix, respectively. Persistent easterly flow due to a strong high-pressure system centered at northern Utah brought monsoon moisture into Arizona. Consequently, convection cells and thunderstorm activities were observed in the northeastern mountains and the southern part of Arizona. Contrary to the June 6th case, cloudiness and micro-scale convective cells confined elevated ozone to the source emission area (central valley) rather than further downwind.

3. Emission Inventory and Processing

3.1 Emission Inventories

Emission inventories were required for both of the Models-3/CMAQ modeling domains, which consists of an "inner domain" with a grid resolution of 2 km x 2 km and an "outer" domain with a grid cell size of 6 km x 6 km to which the inner domain is nested (Figure 3-1).
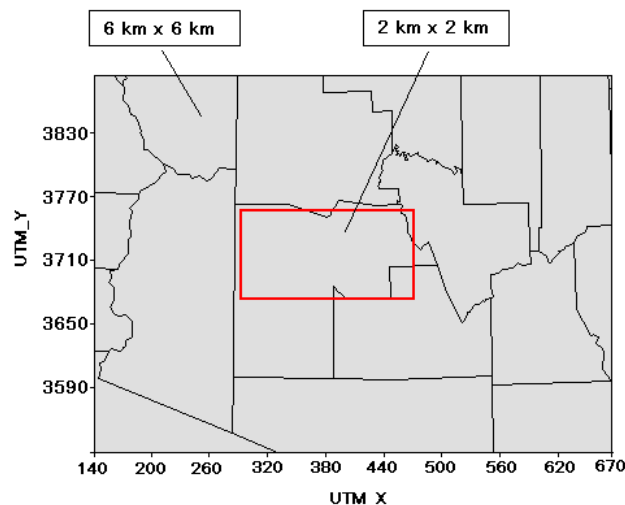

Figure 3-1. Models-3/CMAQ modeling domains with grid resolutions of 2 km x 2 km (inner domain) and 6 km x 6 km (outer domain), respectively.

The 1999 "Ozone Maintenance Plan" emissions inventory was provided by the MAG for the inner modeling domain with a grid resolution of 2 km x 2 km. This 24-hour emissions inventory for a typical summer day contains hourly gridded emissions for the species CO, NO, NO_2, OLE, PAR, TOL, XYL, FORM, ALD_2, ETH, MEOH, ETOH, ISOP, SO_2 and AERO. The Southwest corner of the inventory grid is located at the Universal Transverse Mercator (UTM) coordinates Zone 12,297 km Easting and 3675 km Northing. The extent of the emissions grid is 92 columns and 43 rows.

The emissions inventory for the modeling domain with the 6 km x 6 km grid was processed based on the inventory data of the WRAP base-case scenario 1996 Emissions Inventory. WRAP implemented a regional planning process to provide the necessary technical and policy tools needed by states and tribes to comply with the Clean Air Act goals of protecting the visibility of many national parks and wilderness areas. The regional haze analyses over the western United States is being performed by employing regional scale, three-dimensional air quality models that simulate emissions and their chemical transformations as well as the transport of criteria pollutants and fine particulate matter (PM).

Daily county emissions were used to quantify stationary area sources. Month-specific data for non-road and on-road mobile sources as well as average daily point source emissions were available. Relevant inventory species were CO, NO_x, VOC, SO_2, SO_4, NH_3, PM_{10} and $PM_{2.5}$. All emissions are presented in terms of tons per day per county, except for the point sources where emissions were provided by the location.

3.2 Emissions Processing

The Sparse Matrix Operator Kernel Emissions (SMOKE) modeling system was used to process the WRAP emissions inventory into the formatted emission files required by the Models-3/CMAQ Air Quality Model. SMOKE supports area, mobile, and point source emission processing and also includes biogenic emissions modeling. SMOKE employs the Biogenic Emission Inventory System, version 2 (BEIS2) and version 3 (BEIS3 prototype). The emissions processing used in the present study includes the steps of chemical speciation, temporal allocation and spatial allocation. This means the conversion of pollutant data to chemical species needed for the air quality model, which involves converting spatial-source data from the county to the grid-cell based information and the processing of temporal data with an hourly temporal resolution in a format commensurate with the air quality model.

For the source type specific temporal allocation, WRAP-based temporal profiles and cross-reference profiles for the different source types were applied. The chemical speciation of the inventory species was done according to the Carbon Bond 4 photochemical mechanism leading to thirteen Models-3/CMAQ species – CO, NO, NO2, NR, ALD2, ISOP, TOL, XYL, TERPB, OLE, FORM, ETH, PAR.

The spatial allocation of the mobile and area source emissions to the grid cells of the modeling domain is based on the spatial distribution of the so-called "gridding

surrogate data." This is a dataset developed using the data corresponding to a resolution finer than those used to spatially allocate county emissions to the grid cells. U.S. EPA's 4 km Spatial Surrogate Data set (http://www.epa.gov/ttn/chief/emch/spatial/) covering the entire US was processed using techniques based on Geographical Information System (GIS) technology. The result is a spatial surrogate data file, which contains the fraction of the county surrogate data in each grid cell of the Models-3/CMAQ modeling domain. The surrogates considered are: agricultural and forest areas, airports, land area, housing, major highways, population, railroads, water area, urban and rural areas, urban primary and secondary roads, rural primary and secondary roads as well as urban and rural population. Each emission source type is spatially allocated with a particular type of surrogate data.

The SMOKE model was applied for the periods encompassing the two episodes June 4-7, 2002 and July 10-13, 2002. In addition to the temporal allocation, the hourly plume rise was calculated for the point source emissions based on meteorological data provided by MM5 meteorological model simulations. The emission data of individual sources were merged into gridded hourly emissions. The total daily anthropogenic NO_x and VOC emissions for July 12, 2002 for a part of the modeling domain are shown in Figure 3-2.
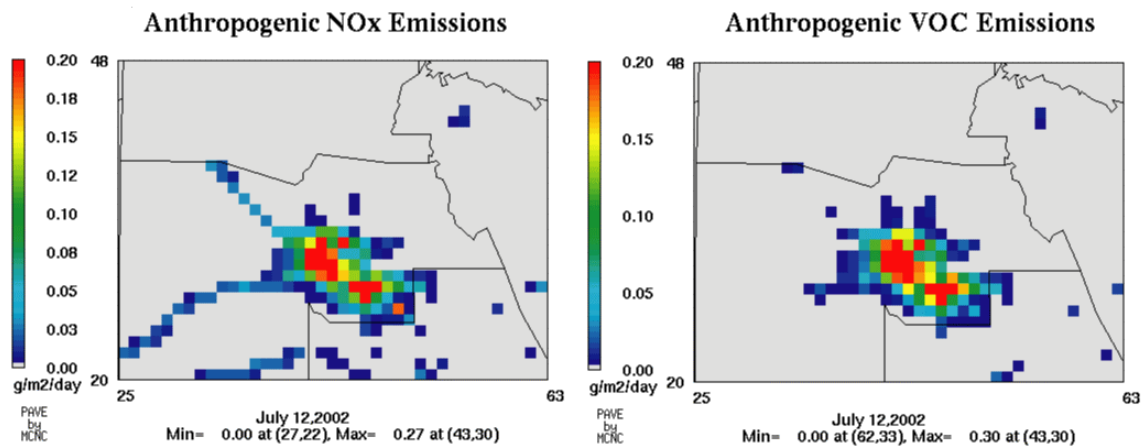

Figure 3-2. Anthropogenic NO_x and VOC emissions for July 12, 2002 as processed by SMOKE based on the WRAP base-case 1996 scenario emissions inventory.

3.3 Biogenic Emissions Modeling

The biogenic emissions were modeled by using SMOKE, which includes a version of the Biogenic Emissions Inventory System 2 (BEIS2) that estimates volatile organic compound (VOC) emissions from vegetation and nitric oxide (NO) emissions from soils. Apart from the land use data, the biogenic emissions depend on the meteorological conditions, in particular the air temperature, incoming solar radiation, wind speed and humidity. Those atmospheric variables were provided for each grid cell

of the modeling domain by the MM5 simulation results. Biogenic emission modeling was carried out for both ozone episodes (Figure 3-3).

Gridded vegetation land use data were prepared using USEPA's Biogenic Emissions Landcover Database (BELD3) that covers the United States, Canada and Mexico with a 1 km x 1 km grid resolution (ftp://ftp.epa.gov/amd/asmd/beld3). Two hundred and thirty land use types are considered in this database. ASU's GIS laboratory helped determine the fraction of each land use type encapsulated in each grid cell of the modeling domain with a 6 km x 6 km spatial resolution.
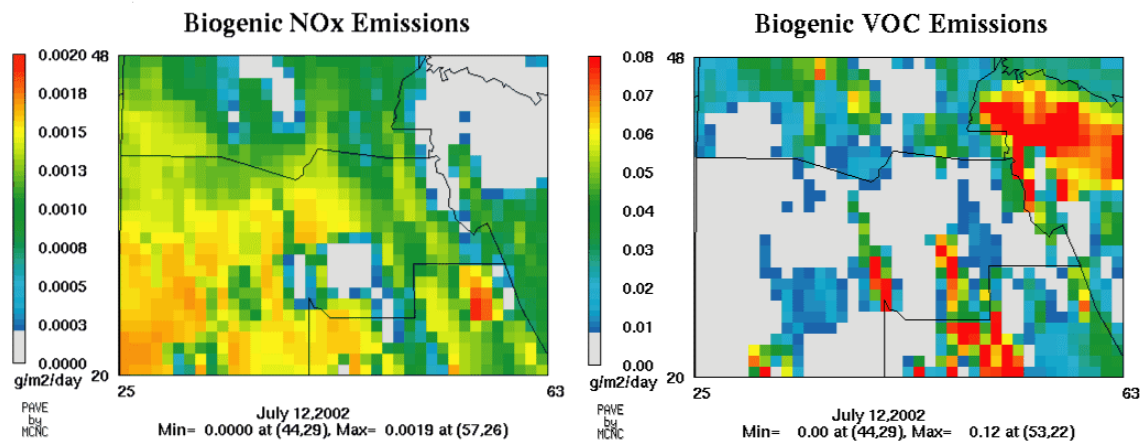
Figure 3-3. Biogenic NO$_x$ and VOC emissions for July 12, 2002 as modeled by SMOKE.

The air-quality model ready inventory data were prepared by merging anthropogenic and biogenic emissions. On-road and off-road mobile sources were the major contributors to the emissions for carbon monoxide, nitrogen oxides and volatile compounds. In order to achieve a higher accuracy for the emissions inventory for the 6 km x 6 km modeling domain, the emissions data of the grid cells, which are in spatial alignment with the extent of the 1999 MAG Ozone Maintenance Plan emissions inventory, were replaced with data from the latter inventory. Those were compiled from their original 2 km x 2 km spatial resolution to 6 km x 6 km grid cells. This procedure is expected to improve the WRAP-derived 1996 inventory for the major source area via incorporating 1999 emissions data.

4. Meteorology Modeling

4.1 Numerical Configuration

The Penn State/NCAR Mesoscale Model, MM5, was employed to provide spatial and temporal distribution of meteorological fields to the air quality model (Models-3/CMAQ). MM5 has been applied to a broad range of studies, including land-sea breeze, mountain-valley circulation, frontogenesis and real-time weather forecasting. The MM5 simulation was performed with 4 nested domains, with respective grid resolutions of 54

km, 18 km, 6 km, and 2 km. The innermost domain spans 224 km x 122 km in E-W and N-S direction, respectively, encompassing the Phoenix valley and surrounding mountains. The 6 km x 6 km grid covers a region 600 km x 420 km in E-W and N-S directions, which is centered at the Phoenix valley. Vertically, 27 layers were used with approximately 10 m agl as the lowest computational layer. The NCEP (National Center for Environmental Prediction) *Eta* model output (Grid 212 with 40 km spacing) was used to provide initial and boundary values for the MM5 simulations and the data assimilation was performed using NWS (National Weather Service) soundings and surface measurements. A period of 67 hours was simulated for each episode: the first 19 hours were considered as the spin-up period, followed by 48 hours of prediction, which included the 24-hour ozone episodes in point and 12 hours of buffer periods fore and aft of the episode.

4.2 Results of Meteorology Simulations

Given that near surface winds are critical for dispersion of pollutants, the analysis was mainly focused on the flow fields. As expected, local thermally driven wind circulation within the valley – up-slope (westerly) flow during day and down-slope (easterly) wind during night – was well simulated by the model. Available wind measurements from ADEQ routine monitoring stations and vertical wind profiles from a Radar Wind Profiler located at the Vehicle Emissions Testing Laboratory site were used to evaluate the model results. Qualitatively, both near surface and upper level winds showed reasonable agreement with the observations. The model performance was evaluated quantitatively using standard statistical tools based on variables such as relative mean bias, mean difference, index of agreement, and RMS vector error. The relative mean bias indicates the fractional difference between the predicted and measured mean to the average of the two. The mean difference is a mean of the difference between the predictions and measurements, whereas the RMS vector error is the RMS of the difference between prediction and measurement of each vector component. The index of agreement and RMSE are measures of the accuracy and error between the predictions and the data; for an ideal model, the former is unity while the latter being zero. Generally, the values of the statistical variables were within the acceptable limits articulated in previous studies: e.g. Pielke and Pearce (1994), Sivacoumar and Thanasekaran (2001), Hanna and Yang (2001), and Lee et al. (2003). These statistical measures are presented in Table 1.

Table 1. MM5 performance measures for surface wind speed and direction for June 6th and July 12th, 2002 cases

	June		July	
	WS	WD	WS	WD
Mean of Obs	2.2	192.4	2.4	171.0
STD of Obs	0.9	83.1	1.3	93.8
Mean	2.1	180.3	2.1	146.6
STD of Obs	1.4	109.5	1.2	123.6
Relative Mean Bias	0.0	0.1	0.1	0.2
Mean Difference	0.1	12.1	0.3	24.3
Index of Agreement	0.5	0.9	0.4	0.8
RMSE	1.5	71.5	1.7	92.7
RMS Vector Error	2.3		3.3	

5 Ozone Modeling

5.1 Numerical Set Up

The Eulerian photochemical model, Models-3/CMAQ (Community Multiscale Air Quality) system, developed by the USEPA was employed to simulate ozone concentrations in the valley and its surrounding areas. Two nested CMAQ domains were used, which are identical to the innermost two domains of MM5, except that several lateral boundary cells were excluded. Observations from ADEQ routine monitoring stations and special measurements during the Phoenix '98 field experiment (i.e. ozone and nitrogen oxides taken by the DOE's G-1 research aircraft as well as hydrocarbon concentrations) were used as initial and lateral boundary values for the outer domain. In order to ameliorate the uncertainty associated with specifying initial conditions, 19 hours of spin-up time was introduced. The selection of a sufficiently large outer domain allowed the typical distances traveled by pollutants by thermal circulation to be smaller than the domain size, thus reducing uncertainties associated with lateral boundary values. The results obtained for the outer domain were used as the initial and boundary values for the inner domain.

5.1 Simulation results.

The monitoring stations were grouped into three categories of West, Central, and Northeast according to their geographic location. Hillside and Palo Verde belong to the West and Pinnacle Peak, Rio Verde, Fountain Hills, Blue Point Bridge, Tonto National Monument, Queen Valley and Humboldt Mountain were classified as the Northeast. The Central category contains Central, South and North Phoenix, Glendale, Maryvale, Surprise, Supersite, South Scottsdale, Tempe, Mesa and Cave Creek. Note that, of the monitoring stations listed above, the Hillside and Tonto National Monument are located outside the inner modeling domain. Therefore, predictions from the outer domain were compared with the observations from those stations while inner domain results were used for the rest of the stations.

Generally, predicted daytime maximum ozone concentration showed fairly good agreement with the observations, while nocturnal ozone concentration showed a deviation from the observations. The nocturnal period, however, is beyond the scope of the present study, which is mainly focused on maximum 8-hour ozone occurring during the daytime.

For the June 6th case, daytime elevated ozone concentration was well captured by the model, except an over-prediction for the western part of the valley (Figs. 5-1,2,3). The elevated ozone in the west of the valley was found to be due to a delay in transition from the nighttime to daytime flows. MM5 simulated persistent southeasterly winds when observation showed a shift from southeasterly to southerly during 1000 – 1200 LST. This prolonged easterly wind transported more ozone and its precursors to the west than in reality. The difficulty of predicting transition is a bane of meteorological models, and the said anomaly points to the necessity of developing accurate parameterizations for transition.

For the July 12th case, the predicted daytime maximum ozone concentrations showed a good agreement with the observations in the west and the central area. The maximum ozone concentration at the Humboldt Mountain was, however, underpredicted (Figs. 5-4,5,6).

When averaged over the 8-hour period (Fig 5-7), the central part of Maricopa county was simulated to be higher than 90 ppb, and its adjacent areas also were found to have elevated ozone > 85 ppb for the June 6th case. The elevated ozone concentration over most of the domain was possibly contributed by the meteorological conditions that were characterized by light wind, clear sky, and deep thermal convection. Conversely, for the July 12th case, the elevated 8-hour ozone was mainly predicted in the vicinity of the Phoenix valley, which was due to limited transport resulting from moist convective cells and thunderstorm activities that were prevalent during that day.
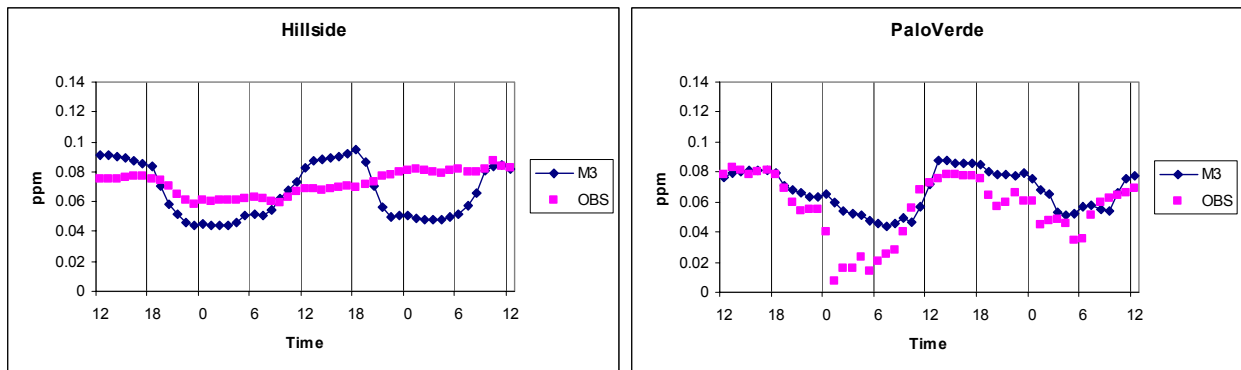

Fig. 5-1. Time series of observed and simulated hourly ozone concentration at stations belong to the 'West' category for the June 6th case.
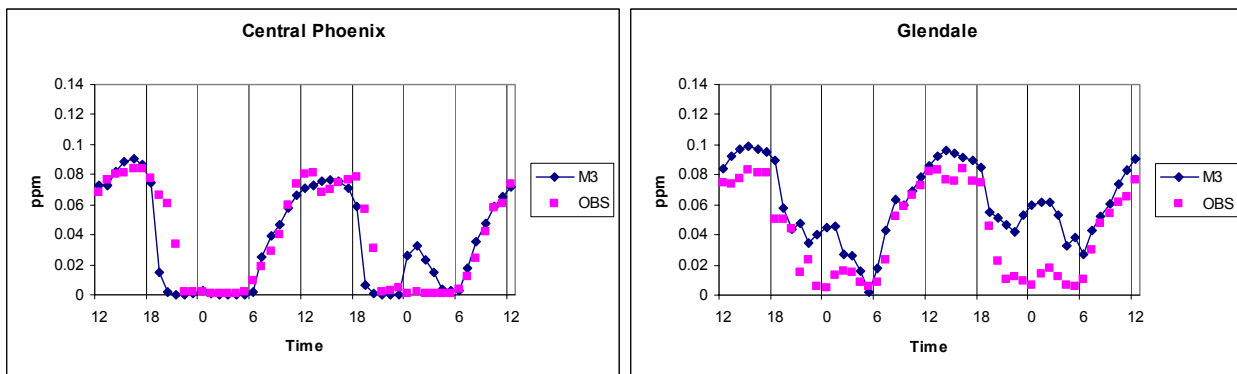
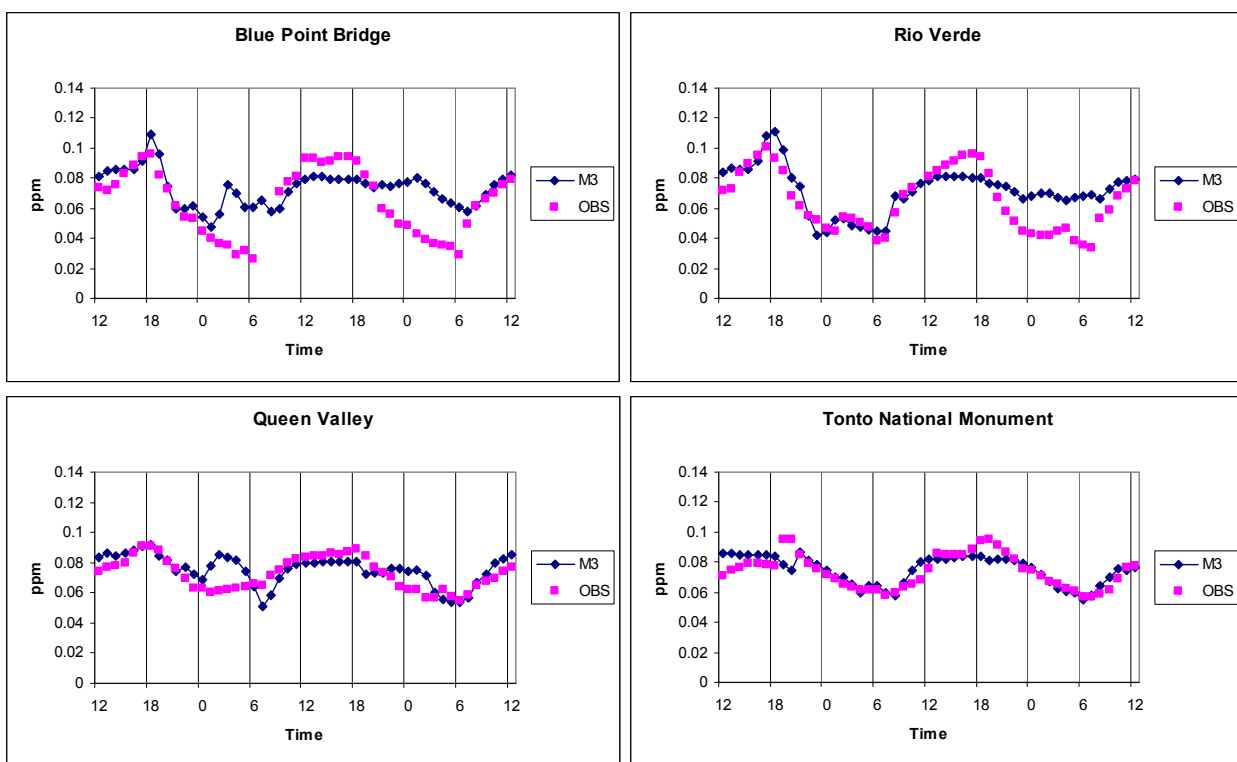Fig. 5-2. Same as Fig. 5-1, except for the 'Central' category.

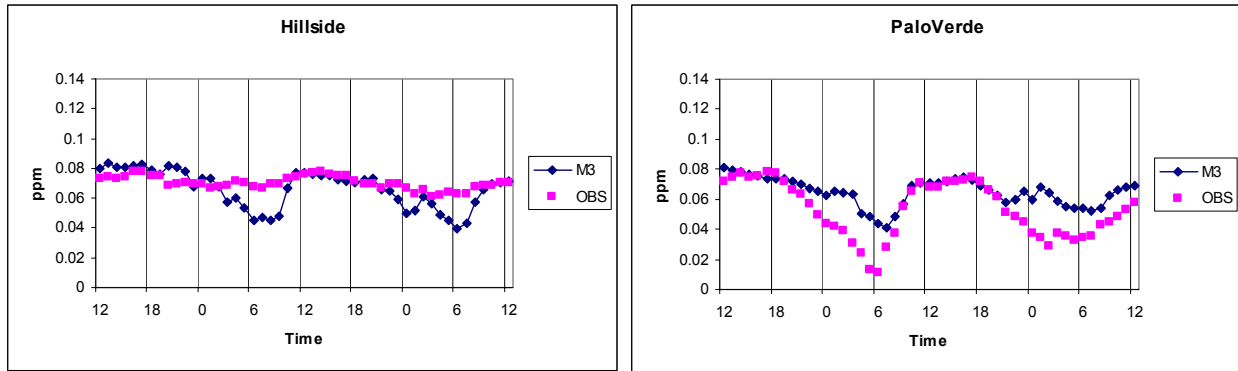Fig. 5-3. Same as Fig. 5-1 except for the 'Northeast' category.

Fig. 5-4. Same as Fig. 5-1, except for the July 12th case.
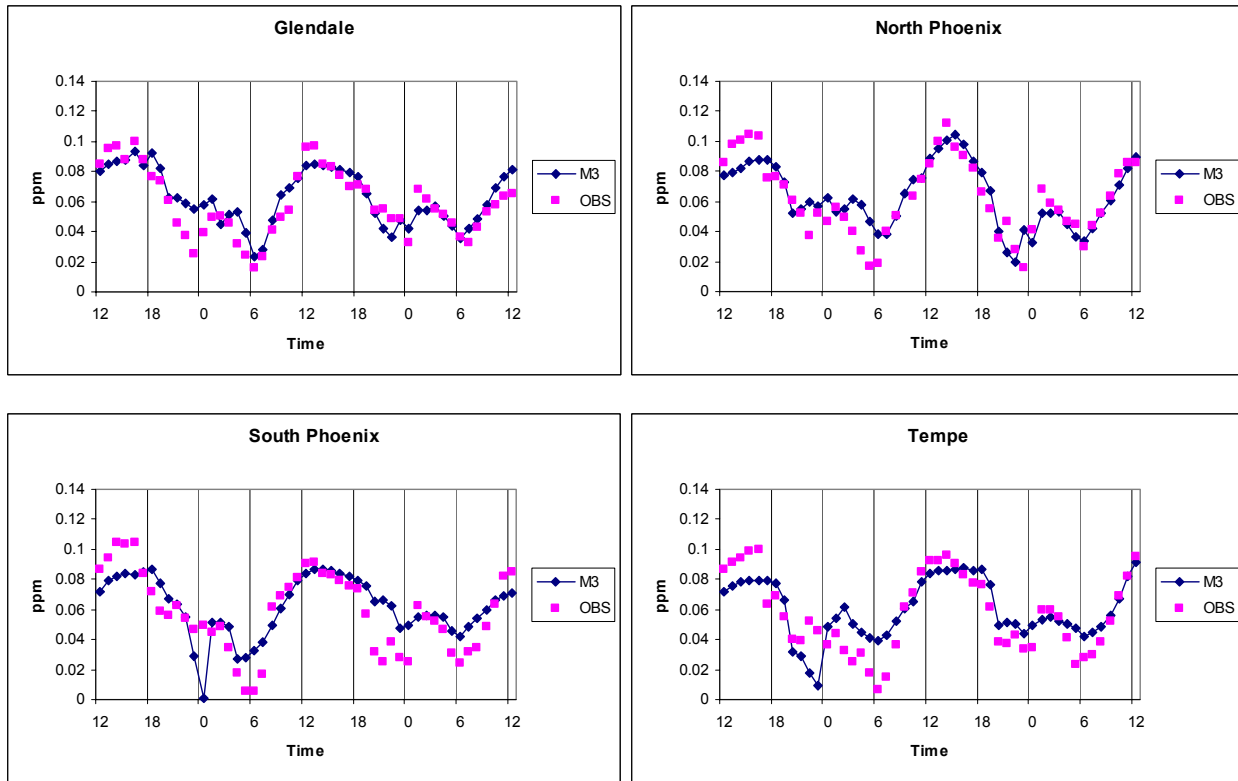

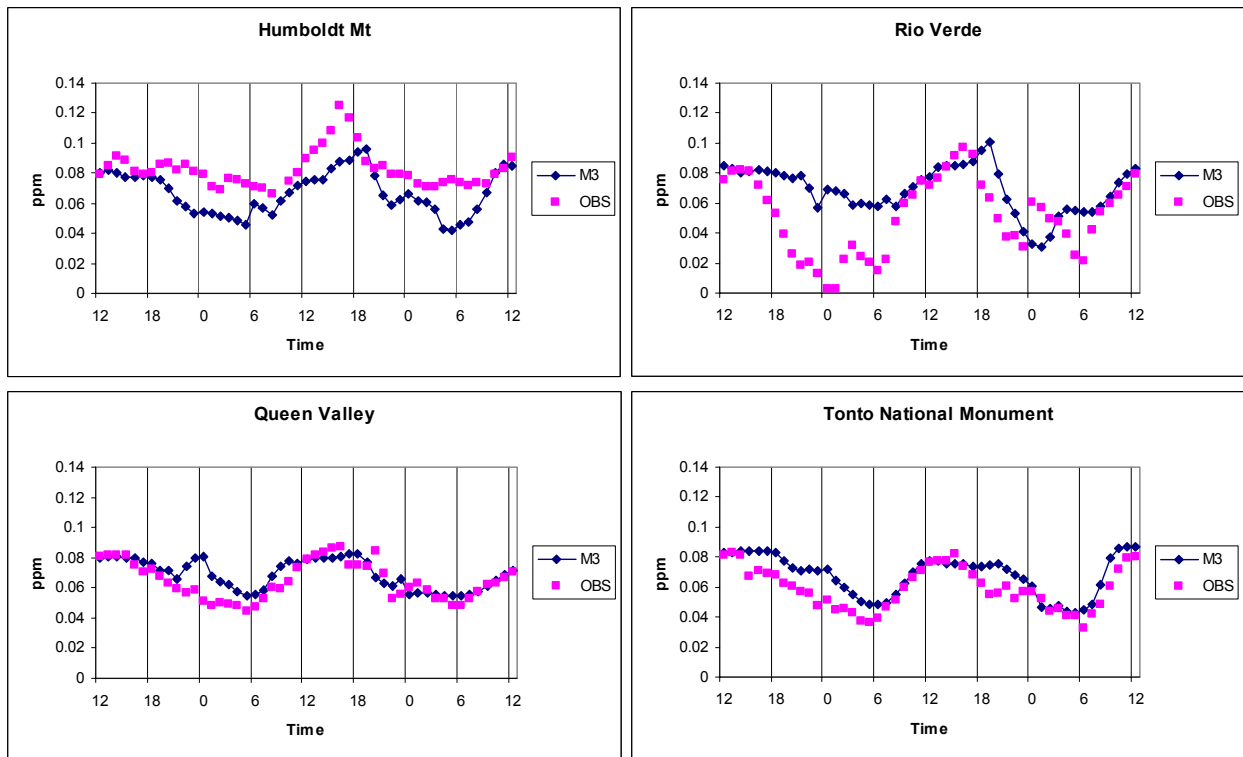Fig. 5-5. Same as Fig. 5-4, except for the 'Central' category.

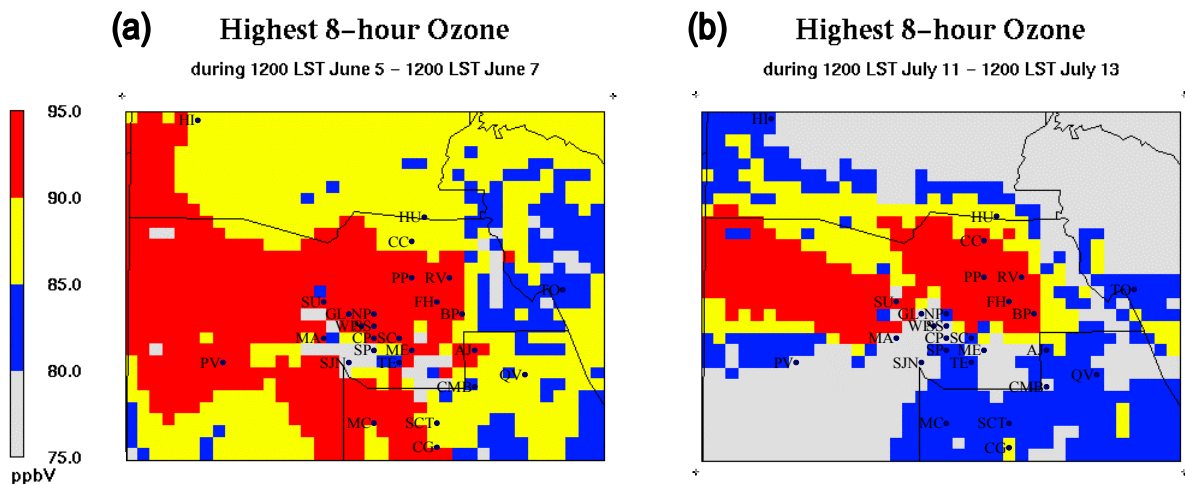Fig. 5-6. Same as Fig. 5-4, except for the 'Northeast' category.

Fig. 5-7. The highest 8-hour averaged ozone concentration during a 48-hour period. (a) June 6th, and (b) July 12th case.

6 Summary

Two design days of elevated 8-hour ozone concentration were simulated by CMAQ, MM5, and SMOKE modeling systems. Two modeling domains were employed: the inner domain is identical to the 1999 MAG Ozone Maintenance Plan emissions domain and the outer domain spans 534 km x 354 km in E-W and N-S direction and is centered on the Phoenix valley. The mesoscale meteorological model MM5 was employed to provide meteorological fields to the CMAQ simulation. Emission inventories for CMAQ are the 1999 MAG Ozone Maintenance Plan and the SMOKE output based on the 1996 WRAP inventory, respectively, for the inner and the outer domains. For each episode, the CMAQ simulation was executed for 69 hours, and the output was analyzed for 48 hours, which encompassed the day of interest and 12 hours ahead and behind of the day.

In general, CMAQ-simulated 1-hour ozone concentration showed a good agreement with the observations for both episodes. For the June 6th case, however, due to the prolonged morning southeasterly flow (or delayed transition) predicted by MM5, CMAQ overpredicted the ozone concentrations in the northwestern part of the valley, and slightly underpredicted those in the northeastern part where the Blue Point Bridge, Rio Verde, Fountain Hill, Humboldt Mountain sites are located. When averaged over an 8-hour period, depending on the meteorological conditions, the central part of the Maricopa County and its immediate surroundings were simulated to have 8-hour ozone concentrations higher than 85 ppb.

References

Hanna, S. R. and R. Yang, 2001: Evaluation of mesoscale models' simulations of near-surface winds, temperature gradients, and mixing depths. *J. Appl. Meteor.* **40**, 1095-1104.

Lee, S. M., J. A. Zehnder, and H. J. S. Fernando, 2003: Evaluation of Meteorological Models, MM5 and HOTMAC, using PAFEX-I data, *J. Appl. Meteor.* In revision.

Pielke, R. A., and R. P. Pearce, 1994: *Mesoscale modeling of the atmosphere, Meteor. Monogr.*, Vol 25, No. 47, Amer. Meteor. Soc., 156pp.

Sivacoumar, R., and K. Thanasekaran, 2001: Comparison and performance evaluation of models used for vehicular pollution prediction, *J. Environmental Engineering*, **127**, 524-530.

Figure list

Table list